

Real-time control system for the Keck Interferometer Nuller: methods and maintenance

Jean I. Garcia, M. Mark Colavita, and Andrew J. Booth

Jet Propulsion Laboratory
Pasadena, CA

ABSTRACT

The real-time control system for the Keck Interferometer Nuller provides the N-band fringe tracking capabilities of the instrument, as well as correcting for atmospheric dispersion in the system. There are three closed-loop servos for controlling the N-band path, as well as two K-band servos which provide open-loop control. A system of synchronized “gates” allows all N-band fringe trackers to operate simultaneously, making it possible to interleave servo corrections with data collection. Several methods of improving servo performance and maintenance of control schemes are discussed.

Keywords: real-time control, control systems, nulling, interferometry, atmospheric dispersion

1. OVERVIEW

The Keck Interferometer Nuller real-time control system¹, responsible for stabilizing the optical path delay between the interferometer arms², requires careful and complex cooperation between a number of subsystems. The system consists of two null-seeking N-band fringe trackers (“nullers”), two corresponding K-band fringe trackers, and a third N-band fringe tracker (the “cross combiner” or XC) which stabilizes the path between the two nullers³. An additional servo loop controls the atmospheric dispersion compensators (ADCs), controlling the dispersion in the system introduced by atmospheric water vapor. Figure 1 shows a high-level overview of the subsystems and how they interact.

1.1 Fringe tracking

The fringe tracking algorithm, utilized by all five fringe trackers of the Keck Nuller, is based on the ZABCD algorithm for phase estimation⁴, in which one delay line is rapidly dithered in a sawtooth pattern over one wavelength. A phase, group delay⁵, and signal-to-noise estimate squared (SNR2) can be calculated from the dithered fringe signal.

The nullers track to an optical path delay corresponding to the null -- the position on the central fringe at which the light from the interferometer arms is interfered destructively. When both primary and secondary nullers are tracking at null, the cross combiner optimizes the path relative to the two nullers, tracking on the fringe signal produced by combining the outputs of the individual nullers. In gated mode (see section 2), the cross combiner servos on the outputs of the nullers at peak.

The K-band fringe trackers help stabilize the N-band path by measuring the error in the common K-N path and sending commands directly to the N-band delay lines. Both the closed-loop servo command and the residual white-light error are sent to the nuller delay lines. Additional commands are sent to compensate for atmospheric dispersion.

Image data for the K-band fringe trackers is provided by a 2-micron camera, which reads out one white-light pixel (for phase estimation) and 5 spectrometer pixels at 2.0-2.4 microns (for group delay estimation). Operating at a frame rate of 200 Hz (that is, one ZABCD image in 5 milliseconds), the K-band fringe trackers utilize an overlapping algorithm which allows them to overlap the current, incomplete frame with

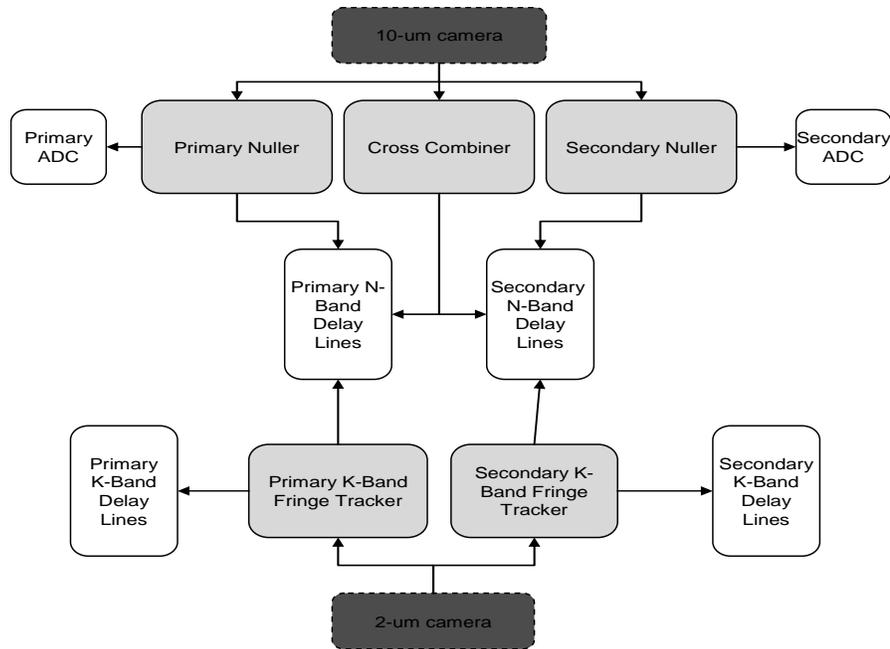


Figure 1. Subsystem overview of the Keck Interferometer Nuller.

data from the previous frame. The result is that the servo can send a correction to the delay line every time a new Z, A, B, C, or D image is acquired, for a servo update rate of 1 kHz.

The N-band fringe trackers, in contrast, do not overlap and therefore require a complete ZABCD frame before a new phase and group delay estimate can be made. The 10-micron camera, which reads out two 48-pixel spectral images (8-12 micron band), acquires a new Z, A, B, C, or D image every 5 milliseconds, for a servo rate of 40 Hz. Because both images are dispersed spectrally, a synthetic white-light pixel is generated and used for phase estimation.

Fringe acquisition occurs as follows: first, the delay lines are commanded to perform a stepped, spiral search, in which each step is approximately 3 wavelengths and the period of each step is the settling time of the group delay filters. After the SNR2 reaches a certain threshold (indicating that the fringe packet has been found), the phase and group delay are used to servo the delay lines to the position which brings the optical path delay to zero. The K-band fringe trackers and the nullers servo directly on their phase estimate, after it has been unwrapped by the group delay. This scheme allows fringe tracking on the higher SNR phase signal and prevents unwrapping errors.

The cross combiner differs from the other fringe trackers in that the dither occurs on internal rapid ramp mirrors whose triangle-wave waveform is provided by a function generator. However, the ZABCD algorithm works similarly, with the exception of a phasor inversion which occurs on the downstroke of the triangle wave. Because it is not limited by SNR2 (tracking on the peak signal results in a 16x boost in SNR2), the cross combiner servos on the group delay estimate. Science data is gathered from the dithered cross combiner fringe signal, at null and at peak.

1.2 Atmospheric dispersion control

In order to achieve a satisfactory null depth, another servo loop corrects for the dispersion introduced by moisture in the atmosphere. The atmospheric dispersion compensator (ADC) consists of two wedges of zinc selenide glass in the optical path, which can be actuated laterally to change the thickness of the glass in

the path. The refractive index of zinc selenide is such that a certain glass thickness can counteract the effect of wet-air dispersion in the path.⁶

The principle behind the ADC loop is: the faster delay line loop locates the position at which the phase estimate over all the pixels is zero. Once this loop has locked, the slower ADC loop is activated, attempting to find a position for the ADC for which the group delay across the pixels is zero. The position at which both the synthetic white light phase and group delay are zero is the position of zero dispersion. Because the phase zero point is arbitrary, this loop assumes that the initial ADC position is within one half wavelength of the global optimum.

The ADC is a relatively slow actuator compared with a fast delay line. The ADC closed-loop bandwidth is less than 1 Hz.

2. THE REAL-TIME SEQUENCE

Because the phase fluctuations on-sky are faster than the null measurement integration time, it is necessary to stabilize all paths during a null measurement. This is straightforward for the K-band fringe trackers and one nuller, because fringe data may be measured simultaneously for all three subsystems. However, this is problematic between the three N-band fringe trackers, because the same detector is used for primary, secondary, and cross fringe data: how does one make it possible to track all three simultaneously?

This problem is addressed by implementing a series of “gates” which track and hold each N-band fringe tracker servo periodically. The real-time sequence consists of eight 25-millisecond beats -- each beat corresponding to one servo cycle of a nuller or cross combiner. From time zero, the sequence is as follows: 1) servo the cross combiner and measure the peak signal for 2 beats, 2) servo the primary nuller for 2 beats, 3) measure the null signal for 2 beats (open-loop control only), 4) servo the secondary nuller for 2 beats. Because each real-time component is tied to absolute time, synchronous operation is straightforward to achieve. There are eleven time gates which are activated at the beginning of a null measurement sequence, each gate tuned in delay so as to eliminate cross-talk. One gate is required for each of the following tasks: 1) track or hold the primary nuller servo, 2) track or hold the secondary nuller servo, 3) move the primary tip-tilt mirrors on or off beam, 4) move the secondary tip-tilt mirrors on or off beam, 5) enable or disable primary delay line dither, 6) enable or disable secondary delay line dither, 7) offset the primary delay line by half a wavelength (i.e. move from the null to the peak of the central fringe), 8) offset the secondary delay line by half a wavelength, 9) track or hold the cross combiner servo, 10) enable or disable primary nuller open loop control, 11) enable or disable secondary nuller open loop control. Figure 2 shows the shape of the white-light signal during gated mode; figure 6 shows the timing waveforms for the various gates.

2.1 Gating the nuller servos

One nuller servo is enabled during 2 of the 8 beats in the sequence. During the 2 beats in which the servo is active, image acquisition and phase estimation is carried out as normal. Fringe data is pushed into its filters, and a position command is sent to its delay line. During the inactive beats, the nuller is held -- although image acquisition is carried out (so that the stream of fringe data is complete), no fringe data is pushed into filter memory, no calculations occur, and no delay line command is sent. In other words, the nuller is perfectly frozen during this time.

2.2 Gating the cross combiner

The cross combiner servo is enabled during 2 of the 8 beats in the sequence, corresponding to the time in the sequence during which the delay lines are positioned at peak. The holding of the servo is identical in behavior as that of the gated nuller servo.

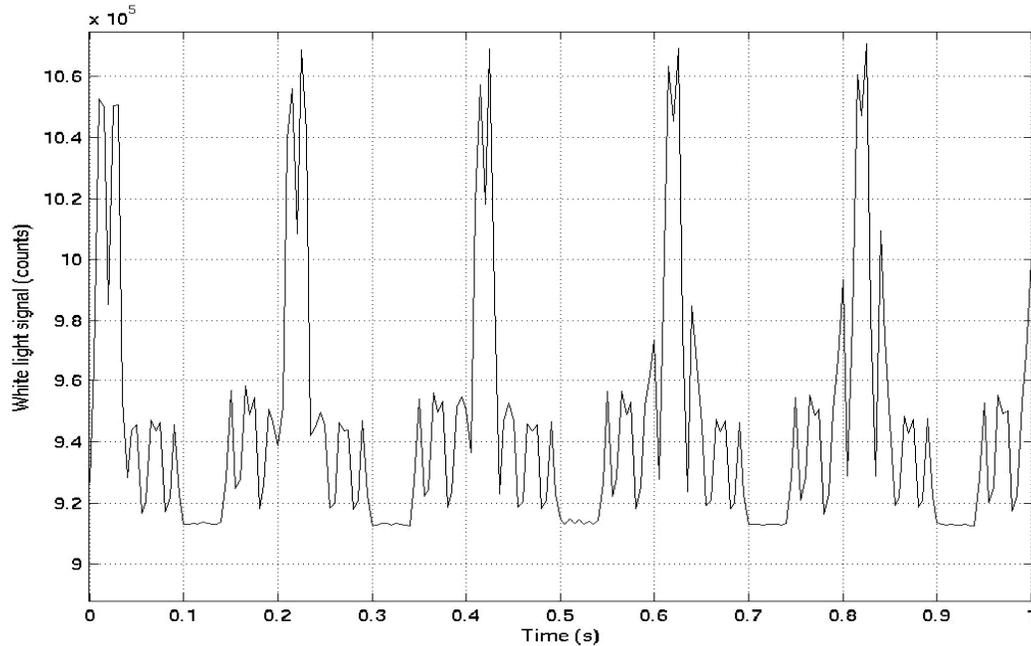


Figure 2. Signal in the synthetic white light pixel, during one second of the real-time sequence. The beats of the sequence are: cross combiner servo active at peak (0 – 50 ms); primary nuller servo active (50 – 100 ms), interference produced by combining nulled outputs (100 – 150 ms), secondary nuller servo active (150 – 200 ms), etc.

2.3 Gating the tip-tilt mirrors

The tip-tilt mirrors must be pointed off-beam for 2 of the 8 beats in the sequence. The primary tip-tilt mirror is pointed off while the secondary nuller is collecting fringe data; likewise, the secondary tip-tilt mirror is pointed off while the primary nuller is collecting fringe data. This gate is implemented by means of a chopping waveform loaded into a function generator, which is triggered by the real-time system. The waveform shape is optimized to minimize rise time while controlling high frequency energy.

2.4 Gating the delay line dither

One delay line's dither is enabled during 2 of the 8 beats in the sequence, or 2 full dither periods. The delay line dither gate is 25 milliseconds behind in phase from its nuller's servo gate. This is so that dithered fringe data will be available to the nuller by the time its servo is activated. The delay line dither is disabled during null measurements, peak measurements, and while the other nuller is collecting fringe data.

2.5 Gating the delay line to peak

For 2 of the 8 beats in the sequence, both delay lines are offset at the same time to the peak of the central fringe. This is accomplished with a simultaneous open-loop command and target update to minimize settling time. During this time, the cross combiner servo collects image data on which to servo.

2.6 Gating the nuller open-loop control

In addition to a closed-loop servo, the nuller performs open-loop correction during the null measurement portion of the sequence. For 4 beats of the sequence, the updated sum of the closed-loop and open-loop commands is sent to the delay line. For 2 beats of the sequence, closed-loop commands only are sent, and for 2 beats of the sequence, the previous closed-loop plus open-loop command is sent. Open-loop control is discussed in further detail in section 3.2.

3. IMPROVING SERVO PERFORMANCE AND SIGNAL-TO-NOISE

3.1 Increasing integration times

The nuller requires a minimum signal-to-noise ratio in order to track. For fainter on-sky sources, special techniques must be used to improve the signal-to-noise ratio.

Most of the filters shown in the nuller dataflow diagram are directly related to the integration time of the servo. (Refer to section 4.1 and figure 5 for information on the nuller dataflow diagram.) A longer filter length corresponds to a longer integration time, and therefore a higher signal-to-noise ratio. For faint on-sky sources, filter lengths are increased by a factor of 10 or 50 from the values used on brighter sources. Several sets of filter lengths are saved and loaded on-the-fly to produce longer or shorter integration times. “Fast” mode, used for internal white-light testing and bright on-sky sources, corresponds to a 50 millisecond phasor integration time. “Medium” mode, the default mode for on-sky observing, has a 500 millisecond integration time, but a 10x improvement in signal-to-noise squared. “Slow” mode, 50 times slower than fast mode, has a 2.5 second coherent integration time and a 50x improvement in signal-to-noise squared.

Although the improvement in signal-to-noise is necessary to make observation of faint sources possible, the drawback to using longer integration times is a significant hit in servo bandwidth. Fringe search integration times must be increased in order to accommodate the longer filters; the servo requires significant lengths of time to settle (up to a few minutes in slow mode); this, combined with the 4x bandwidth hit from servoing in gated mode, and it becomes clear that several methods of open-loop control must be utilized.

3.2 Open-loop control

The most important open-loop control system comes from the K-band fringe tracker. The K-band fringe tracker has a considerable advantage in servo bandwidth, computing a new servo error once every millisecond. (Compare this to the nuller in gated mode, which sends a delay line command only once every 200 milliseconds, *and* is slowed down by an additional factor of 10 or 50 in medium and slow modes!) Also, the K-band fringe tracker, operating at a shorter waveband, is inherently more stable in phase than the N-band fringe tracker. The K-band fringe tracker is capable of correcting a significant amount of error in the common K-band and N-band paths.

The K-band fringe tracker sends open-loop commands directly to the nuller delay lines. Therefore, the nuller control system is unaware of and isolated from the open-loop feed forward from the K-band fringe tracker. The nuller servo only sees that its phase measurement is more stable when feed forward is enabled. The K-band fringe tracker sends its own high-bandwidth servo command to the nuller delay lines, as well as its residual error. In addition to this dry-air feed forward, the K-band fringe tracker also computes dispersion control feed-forward components which are sent to the nuller delay lines and ADCs.

Because the science data occurs in the gated sequence when closed-loop control is disabled, the nuller also employs internal feed-forward control when gating is enabled. During the beats of the gated sequence when the servo is active, the nuller delay line target consists of the feedback command only, preventing servo instability. During the remaining 150 milliseconds, including the time during which science data is collected, the closed-loop servo is inactive. At this time, a feed-forward command consisting of the sum of the feedback command and a fraction of the residual servo error is sent to the delay line. The step response with internal feed forward is shown in Figure 3.

Another method of compensating for low servo bandwidth is the “big step.” After N-band fringes are first acquired (but before closed-loop control begins), a delay line target corresponding to the group delay estimate at the time of fringe acquisition is sent. This brings the delay line close to zero group delay in one settling-time period. With a medium-mode or slow-mode servo, the big-step allows the servo to settle in a significantly shorter period of time.

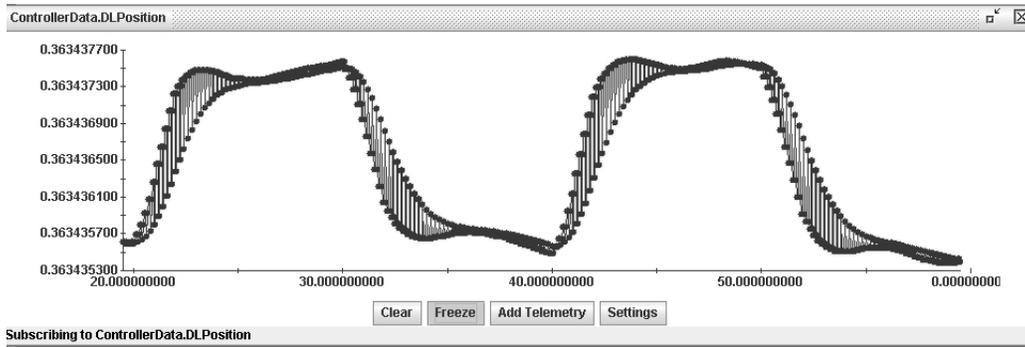


Figure 3. The commanded delay position, tracking out a square wave modulation. The slower waveform is the closed-loop control; the faster waveform is the open-loop feed forward.

4. MAINTENANCE

4.1 Dataflow diagram

The K and N-band fringe trackers are encapsulated by a dataflow diagram (figures 4 and 5 below), which concisely describes the computational chain, beginning at image acquisition, through phase and group delay estimation, and ending at the final delay line command. Every filter and gain is included in the diagram, as well as describing important concepts in the dataflow, such as switches, limits, caches, and thresholds. The dataflow diagram provides an indispensable tool for facilitating modification of and communication about the real-time algorithms. Key telemetry items are also noted, showing where in the data chain the relevant telemetry is pushed.

The dataflow diagram has the ability to capture parallel paths of computation in the real-time control algorithms. There are several filters which must be of the same length in order to maintain consistent data age and latency. Also, the dataflow diagram notes which filters are key in increasing the coherent integration time to improve the signal-to-noise ratio.

4.2 Timing diagram

This diagram, figure 6 below, assists in managing the many gates utilized by the system. With each new feature added for gated mode, this diagram becomes very important for organizing and recalling the complex timing scheme of the system.

4.3 “Go-slow” table

Another table, the “go-slow” parameter set, is used for identifying which filters need to be lengthened, the lengths of the filters, their names in the configuration database, and the settling time and servo bandwidth of each mode. This table is useful for managing parameters, noting servo rise times and comparing past performance to present performance, and listing which features are enabled and disabled to provide basic nulling functionality.

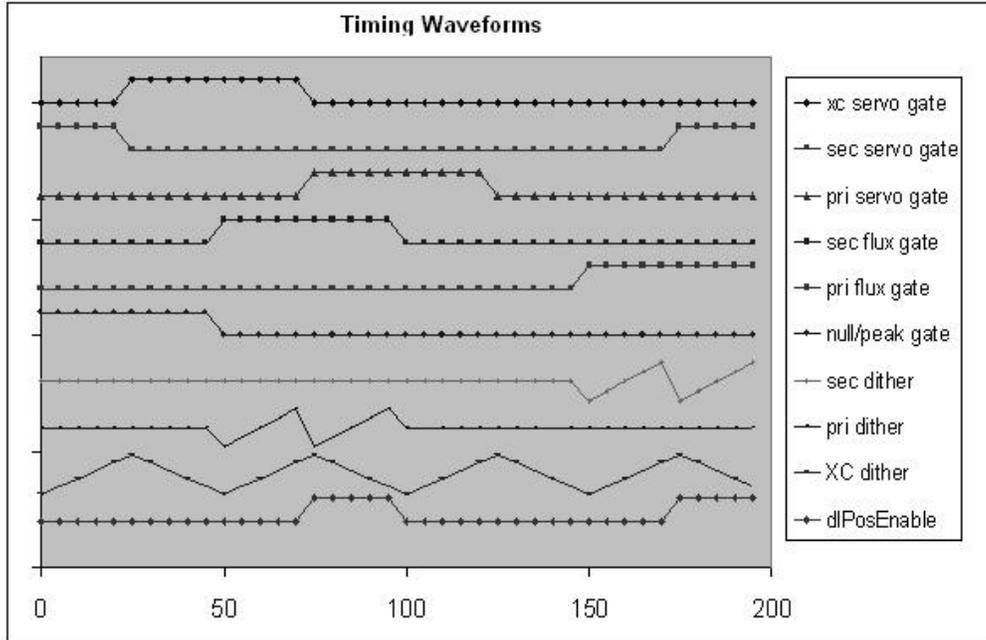


Figure 6. The timing waveforms for achieving simultaneous, gated control. X-axis is time in milliseconds. Each graph tick is 5 milliseconds; each servo cycle, dither sweep, and gated beat is 25 milliseconds.

5. CONCLUSION

When the idea was first conceived to have a real-time system which consisted of five fringe trackers, each with its own feed-back and feed-forward paths, it was clear that it would prove to be an ambitious endeavor. Since then, techniques for providing simultaneous tracking and increasing SNR² have been developed, each allowing greater flexibility to the system but hindering servo performance, thus necessitating the implementation of new methods of open loop control. These added complications emphasize the importance of maintaining a proper grasp of logical and algorithmic implementations of the system (in this case, by means of detailed documentation). Adherence to this principle is of immense benefit, helping the developer keep the control system “under control.”

ACKNOWLEDGEMENTS

The work presented here was conducted at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration, and at the W.M. Keck Observatory, California Association for Research in Astronomy. Funding for the Keck Interferometer project is provided by the National Aeronautics and Space Administration.

REFERENCES

1. A.J. Booth, M.M. Colavita, J.I.Garcia, C.D. Koresko, “The control system for the Keck Interferometer nuller”, these proceedings.
2. M.M. Colavita, P.L. Wizinowich, and R.L. Akeson, "Keck Interferometer status and plans", *Proc. SPIE* Vol. 5491, *New Frontiers in Stellar Interferometry*, Wesley A. Traub Ed., p. 454-463, 2004.

3. E. Serabyn, et al., "The Keck Interferometer Nuller (KIN): configuration, measurement approach, and first results", *Proc. SPIE* Vol. 5905, *Techniques and Instrumentation for Detection of Exoplanets II*, Daniel R. Coulter Ed., p. 272-283, 2004.
4. M.M. Colavita, "Fringe Visibility Estimators for the Palomar Testbed Interferometer", *Publications of the Astronomical Society of the Pacific*, **111**, p. 111-117, 1999.
5. P.R. Lawson, M.M. Colavita, P.J. Dumont, and B.F. Lane, "Least-squares estimation and group delay in astrometric interferometers", *Proc. SPIE* Vol. 4006, *Interferometry in Optical Astronomy*, p. 397-406, Pierre J. Lena and Andreas Quirrenbach Eds., 2000.
6. C.D. Koresko, B.P. Mennesson, E. Serabyn, M.M. Colavita, R.L. Akeson, M.R. Swain, "Longitudinal dispersion control for the Keck interferometer nuller", *Proc. SPIE* Vol. 4838, *Interferometry for Optical Astronomy II*, Wesley A. Traub Ed., p. 625-635, 2003.